Pro Gradu
Economics

Effects of Agglomeration on Labor Productivity – Empirical Cross-country Evidence from the Nordics

Jaakko Järvinen

April 2019

Advisor: Prof. Markku Lanne

UNIVERSITY OF HELSINKI
Faculty of Social Sciences

PL 54 (Unioninkatu 37)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Social Sciences	Master's Programme in Economics

Tekijä — Författare — Author

Jaakko Järvinen

Työn nimi — Arbetets titel — Title

Effects of Agglomeration on Labor Productivity – Empirical Cross-country Evidence from the Nordics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro Gradu	April 2019	50 pages

Tiivistelmä — Referat — Abstract

The main thesis of economies of agglomeration is that by increasing the density of employment, economic benefits will follow. In this research, this hypothesis is tested in the context of the Nordic countries by studying if the increases in the employment density affect the regional productivity. This effect between the employment density and regional productivity is called the agglomeration effect.

The theoretical background of this effect lies in three fundamental concepts: economies of scale, labor pooling and knowledge spillovers. Cities have their origins in the economies of scale and agglomerations of people they comprise of form a fertile base for effective matching between employers and employees. The denser these production centers are populated, the easier it is for the spillovers of innovation and ideas to happen.

This study uses a linear ordinary least squares (OLS) model to estimate this effect. The data consists of 70 regional observations and the model comprises of employment density as the explanatory variable, varying number of education level control variables and dummy variables for different countries. Endogeneity of the explanatory variable is also assessed but as the proposed instrument, the total land area of the included regions, proves to be invalid for this particular geographic region, the OLS estimates will serve as the final results.

In the previous studies conducted in Europe and in the USA, the magnitude of the agglomeration effect has been found to be between 4.4 and 6 %. This study's estimates tell the effect to be between 2.1 and 2.9 % in the Nordic countries that is lower than the corresponding values for the aforementioned regions. This result is discussed to stem from the unique geographical and political characteristics of the Nordic regions.

Avainsanat — Nyckelord — Keywords

Agglomeration, Population density, Productivity, Nordic countries

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

Contents

1. Introduction

In the field of agglomeration economies the consensus is that when people and firms locate near one another benefits can be achieved (Marshall 1920; Henderson 2003a; Rosenthal and Strange 2004). Main source of these benefits is the transportation cost savings that mean not only costs for the exchange of goods and people but also ideas (Duranton and Puga 2004). During the last century, transportation costs have fallen dramatically but still people and firms tend to agglomerate close to each other. This kind of behavior can be seen by studying the current megatrends that are taking place all around the world; people pack themselves into cities and companies of same fields form big clusters.

It is easy to see the benefits that stem from this kind of activity. In the case of modern knowledge-based work, most of the transportation costs come from transporting workforce as workers are the main source that create added value (Stewart and Ruckdeschel 1998). If some location already has companies of the same field, there will probably also be workforce living there. In this sense, a firm benefits from locating itself to that particular location. The same is also true from the workers' perspective: a city with a concentration of industries is probably also a city with a lot of amenities. Modern people rely on services and by basing their location choices on existing agglomerations many benefits may come by (Baum-Snow and Hartley 2016).

The main thesis of economies of agglomeration is that benefits brought by it outweigh the costs. In the literature, main benefits are often identified as increased labor productivity (Ciccone 2002; Ciccone and Hall 1996) or even significant increases in the GDP growth of a country (Brülhart and Sbergami 2009). On the other hand, an uncontrollable agglomeration might bring some major disadvantages. These disadvantages can materialize, for example, in the form of congestion that is happening in many Asian and African megacities at the moment, or rising prices of housing. Like the theory suggests, it would be easy to conclude that designing and building as dense and populous cities as possible would be the answer to boost the labor productivity or GDP to unthinkable levels but, luckily, the answer is not that simple.

A common case in economics is that there are often many driving forces behind one seemingly simple phenomenon. Modern econometric techniques and the vast collection of relevant data has made the analysis of these forces easier but the part that remains difficult is the identification of the root causes. As might be evident, thorough econometric analysis is not enough to successfully do this. A certain level of knowledge of economics in general and some historical background of the given geographical region are essential. When these competences are added together with the theoretical work offered by the existing literature, it is possible to gain some understanding of these causes acting in the background.

In this study, the effect behind all this is assumed to be an effect called *agglomeration effect*. This agglomeration effect can be studied by inspecting economic, geographic and demographic indicators in a fine geographical level and deriving results from the analysis. These results would then provide invaluable information about the mechanics between location choices of individuals and the economic productivity of a country.

1.1 Research question

The main research question of this study is to find out if there is a relationship between the agglomeration of people and the labor productivity across the NUTS 3[1] regions of the Nordic countries (Finland, Sweden, Norway and Denmark). Iceland is being left out since there exists only two Nuts 3 regions in the country: Reykjavik area and the rest of the country and the existence of only two regions does not allow to conduct a proper econometric analysis.

In addition to finding out the existence of the effect, this study aims to finding out the magnitude of the effect. It will be interesting to see how the distinct geographical, political and demographical characteristics of the Nordic countries affect

[1]NUTS stands for the nomenclature of territorial units for statistics and is a geographical dividing system for the European Union

the magnitude of this effect. Based on the economic theory, an initial prediction is that this effect would be smaller than, for example, in the USA and larger European countries. Reason for this might be the Nordic countries' lower levels of urbanization and a remote location to the European core market.

The main motivation for this research is to fill the existing gap in literature and to gain valuable insight that can be used as a base for political decision making concerning regional population issues. To further motivate the importance of this research, a brief look on the ongoing public debate can be taken. The disparities between economic development levels between centers of growth and the countryside is a particularly current topic at least in the Finnish politics as parties are currently updating their stances on the issue for the next electoral term. Historically, there has been an impressive continuance of political will to maintain a large scale income redistribution within more and less productive regions and also many services and benefits have been granted by law for every citizen, no matter where they live. This kind of public policy has been in effect for decades, until the most recent 2015 government have gotten rid of some of these obligations.

This connects nicely with the research question of this study as the obtained results might be of use to further improve the argumentation of this public debate. In order to make decisions concerning the scale of these supportive measures, it is essential to know how much more productive the centers of growth really are. Among many economic insights, this political aspect is surely one of the most interesting results this empirical research will provide.

When it comes to the challenges in the analysis, Nordic countries cannot, in any sense, be described as similar to other European countries. Vast geographic areas, sparse population across the regions and some historical policies make them a challenging group to compare to the other countries in same development levels. The fact that the urbanization of population started not until in the 1970s in Finland and a couple of decades earlier in the other countries tells that these countries' largest cities have a completely different history than their peers in Europe and in the USA. In addition to this, the geographical isolation have surely contributed to the economic development of the Nordic countries and made outside influences more rare. All this might derive some interesting insights and therefore the results of this study and the results for other countries should not be compared without a fair amount of criticism.

2. Literature review

The literature on the economies of agglomeration dates back to the 1920s. However, topics of industrial organization and location choices have been of a great interest for economists for a far longer time. Since the fundamental studies of Adam Smith (1776), the forces behind the location of companies have stirred some of the most intense academic discussion and debate concerning mostly the causality issue and reasons associated with these choices.

The roots of agglomeration as a phenomenon lie in the economical specialization of industries that can be considered to have started after the first Industrial Revolution. During the 1800s first specialized fields emerged and the technological innovations allowed the rise of modern textile and manufacturing industries. The current economic division of labor between different countries would not be possible without these events and the technological developments during this time contributed essentially to the incredible rise in the standard of living all around the Western world.

Since David Ricardo's (1817) famous theory of comparative advance, it has been widely accepted that regional specialization is the key to prosperity. Back then, firms' location choices were mostly based on the existence of natural resources and trade relations between empires played a crucial role in the division of wellbeing through foreign goods. In the modern day information-based work it is not anymore about the natural resources but labor resources; in many fields skilled working population forms the most important part of companies' assets and its location choices have to be taken into account when deciding the optimal location for a firm. This is where economies of agglomeration come to the picture.

The most important concepts of economies of agglomeration will be explained in the following four sections. The underlying theoretical model will be discussed in the light of the existing literature in the fifth section and the most relevant empirical results will be assessed in the last section of this chapter. The model specifications in the latter part of the thesis are based on the publications presented in this chapter.

2.1 Economies of scale

As Marshall (1920) found out in his field-defining publication, there are three main sources through which economies of agglomeration affect the economy: pooling of skilled workforce, knowledge spillovers and local linkages between producers and suppliers. All of these have an important role in explaining the effect of agglomeration and therefore have to be addressed together. The fundamental force that all of this bases on is the economies of scale.

In short, modern cities have their foundations in this economic phenomenon. When more people are available within a reasonable distance, the potential market for products also grows faster than the transportation costs and also the labor pool increases. For a firm, this makes same or even higher benefits available with less costs. As discussed earlier, in order for the firm to benefit from the economies of scale (e.g. having an increasing factor productivity) the benefits of agglomeration have to outweigh the costs. This can be seen as one of the fundamental assumptions concerning the economies of agglomeration.

To this interpretation bases also one important theory developed by the world-renowned economist Paul Krugman (1991): New Economic Theory. As an explicative theory for agglomeration, it is in close connection with benefits of agglomeration and therefore an important part in understanding the reasons of economies of agglomeration. In New Economic Theory, Krugman analyzes agglomeration from the economic geography's point of view and bases his arguments on the economies of scale approach; if some region already has high productivity in some field, the other companies will also relocate into that region. In this kind of situation the causation can be interpreted as running from productivity to agglomeration. This can be seen as a challenging view towards the traditional theory on the effects of agglomeration.

2.2 Labor pooling

An important observation by Marshall that follows from the localized agglomeration of inputs and companies of production is the positive effect of the proximity of firms on labor force. This allows the process of finding proper workforce to be more efficient and also makes the recruitment cheaper. An optimal location between the companies and labor force allows matching employers and skilled, suitable employees to be far easier than in the case of regional remoteness. A good illustration of the phenomenon is the study conducted by Arzaghi and Henderson (2008) in which they analyzed Manhattan-based advertising firms and the benefits experienced from networking with competitors of same field located close by. Their main finding was the rapid spatial decay of benefits since most of the benefits were achieved by being an active member of the community.

These observations by Marshall and Arzaghi and Henderson shed some light on another phenomenon which is called the reverse Williamson hypothesis (Williamson 1965). According to it, the existing clusterization of well performing companies in a region attract more companies to share the benefits. The causation in this case can be seen to run bidirectionally since the agglomeration itself surely increases the productivity but also the productivity gains and success experienced in a region work as the main force of attraction. This is certainly a phenomenon requiring further research.

2.3 Knowledge spillovers

Marshall also made observations how positive spillovers happened through agglomeration and one of them considered *informal* information spillovers. These spillovers were considered as unintentional happenings during which information was passed along between workers of the same region.

A good example of this phenomenon is a study of educational spillovers' effect on productivity (Moretti 2004) in which a city where the fraction of college graduates grew faster was found to experience higher productivity than a city where the fraction grew slower. The effect was found to be largest within companies operating in the field of high-technology and was also proven to be due to the interactions between companies from different fields. This supports urbanization as the explaining phenomenon for increased productivity.

2.4 Localization and urbanization economies

In addition to these, Henderson (2003b; 2003a) has investigated the presence of localization and urbanization economies. Concepts of localization and urbanization economies mean that companies within a certain industry (localization economies) or of different industries (urbanization economies) experience agglomeration as a benefit. While concluding that urbanization has a major effect on the productivity growth of the cities, he makes an important observation: there exists a certain level of favoring between cities. The national capital may be favored in the advantage of politicians and some of the smaller growth centers might be neglected completely. Therefore, the share of people living in the largest city and the possibilities of agglomeration might depend heavily on national policies and institutions.

Same effect is also taken into account by Quigley (2008) who states that it seems to be a problem mainly in developing and politically less democratic countries. Ways of favoring urban populations might be unequal price setting between urban and rural products, uneven infrastructure investments or even direct legislative actions.

Good examples of government intervention are China's migration policies in which rural population's movement from the rural areas is being restricted since it is believed to have a negative effect on the economic growth of the countryside. This effect has been studied by Au and Henderson (2006) and it might have a surprisingly large impact on the agglomeration effects in the developing nations.

This fairly remote example connects with this study in the form of political similarities between China and the Nordic countries. Each of the Nordic countries has a party driving the rights of the agricultural workers and people living outside the main cities in sparsely populated areas. This political influence is anything but insignificant; parties representing this ideology are currently in the position of the leading party in Finland and the fourth largest in Sweden and in Norway.

Traditional policies driven by these parties include opposing many essential infrastructure investments in the cities and fighting to maintain the highly subsidized agricultural industry in the places not necessarily suitable for it. Of course, investments aimed to sparsely populated areas are necessary but observing critically from the economic perspective these kind of policies can be seen to cause more growth reducing than growth promoting effects.

The effect this kind of political intervention has on the economic development of the most important centers of growth is hard to estimate precisely but the amount of subsidies flowing to these regions is anything but insignificant. This might be one of the factors causing differences between the agglomeration effects in the Nordics and in the studies conducted in the United States and European countries.

2.5 Theoretical background

The theoretical framework to study the agglomeration effect across countries that is used as the basis of this thesis has been developed by economists Ciccone and Hall (1996) in their publication studying the agglomeration effects in the US. In this study, they form two models that both regress density at the country level to the labor productivity at the state level. By density they mean the employment, human capital (education) and physical capital density in a given region.

In this study, Ciccone and Hall make two alternative assumptions and use two different techniques: nonlinear least squares and nonlinear instrumental variables regression. In the first case, they assume the productivity (dependent variable) to be uncorrelated with the explanatory variables but admit that this identification might not actually produce reliable estimates due to the large measurement error in productivity across states. The other case would be that the employment density would be somehow correlated with the differences in the states' productivity that might be caused by climatic or geographic features. This seems highly intuitive as US' states differ highly in their characteristics and location.

In the instrumental variables (IV) regression, they use four different instruments: proximity to the eastern seaboard, existence of railroad in 1860, population in 1850 and population density in 1880 and justify the usage of these instruments by stating that they have contributed to the legacy of modern agglomerations but have not actually affected the productivity we are observing now.

In another study conducted a few years later, Ciccone (2002) developed the model further and, in turn, studied the agglomeration effects in the five largest European countries (by population). Additions he has made to the model are the inclusion of regional and country dummies and a slight modification to the education parameter: it now allows the usage of education levels more flexibly i.e. the amount of possible education levels do not have to be constant across regions. The estimating equation the model yields is the following:

$$log\frac{Q_{sc}}{N_{sc}} = \theta(log\frac{N_{sc}}{A_{sc}}) + \sum_{e=1}^{E_c} \delta_e F_{esc} + \text{Country/regional dummies} + u_{sc}, \tag{2.1}$$

where $log\frac{N_{sc}}{A_{sc}}$ is the employment density in a region, F_{esc} the education level of workers in a region, u_{sc} the error term and $log\frac{Q_{sc}}{N_{sc}}$ the average labor productivity in a region. Country/regional dummies are included to control for the differences in exogenous total factor productivity and rental prices of capital between countries and regions.

Like the previous study, also this is conducted using two different approaches in estimation: ordinary least squares (OLS) and IV regression. It is stated by Ciccone that the OLS approach would not produce consistent estimates in case regional fixed effects do not efficiently capture exogenous differences in total factor productivity across NUTS 3 regions.

This possibility of endogeneity is taken into account by using an instrumental variables regression approach. In the paper, Ciccone resorts to using the total land area of NUTS 3 regions as the instrument and justifies the decision with history: the borders of NUTS 3 regions have been determined hundreds of years ago and therefore can not affect the modern productivity, albeit, they are negatively correlated with the modern employment densities in regions. These two characteristics fulfill the requirements of a valid instrument.

2.5.1 Model formulation

This subsection formulates the model used in the Ciccone (2002) paper. The objective is to gain an understanding of the components affecting the final model and also the rationale behind the theory. The formulation is presented in a pretty fine detail so it might be good to keep in mind at this point that, as seen in the previous section, the estimating equation actually turns out to be relatively simple.

Let's start the formulation of the model by defining the production function. This production function depicts the production on a single square kilometer of land in a region, distinguished by subscript sc, and is defined as a function ($f()$) of five parameters multiplied by the total factor productivity Ω_{sc}. The following notation will be used throughout the rest of the thesis: all parameters with subscript sc depict regional values for a variable and parameters with no subscript are measured in the level of Nordic countries. Let's define the production function as follows:

$$q = \Omega_{sc} f(nH, k; Q_{sc}, A_{sc}). \tag{2.2}$$

As can be seen, the output q in a square kilometer of land is specified as a function of the number of employed people n, the average level of human capital of the workers H and the amount of physical capital used k. Variable Ω_{sc} represents the index of region's total factor productivity , Q_{sc} total production in a region and A_{sc} total area of a region.

Assuming the elasticity of output-per-square kilometer with respect to the regional density of production to be constant, we'll get the following form for the production function:

$$q = \Omega_{sc} f(nH, k; Q_{sc}, A_{sc}) = \Omega_{sc}((nH)^{\beta} k^{1-\beta})^{\alpha} \left(\frac{Q_{sc}}{A_{sc}} \right)^{\frac{\lambda-1}{\lambda}}, \tag{2.3}$$

where α depicts returns to capital and labor on the square kilometer of land, λ explains the externalities arising from the density of production and β parametrizes distributional characteristics that will be assessed later.

As our eventual objective is to form an equation for the dependent variable of the model, the average labor productivity, we proceed by forming a function depicting the aggregate production. Marking the aggregate production in each region by Q_{sc}, we multiply the production function (2.3) by the total area of regions A_{sc}:

$$Q_{sc} = A_{sc} q = A_{sc} \Omega_{sc} \left(\left(\frac{N_{sc} H_{sc}}{A_{sc}} \right)^{\beta} \left(\frac{K_{sc}}{A_{sc}} \right)^{1-\beta} \right)^{\alpha} \left(\frac{Q_{sc}}{A_{sc}} \right)^{\frac{\lambda-1}{\lambda}}. \tag{2.4}$$

Here, N_{sc} is the total employment in the region, H_{sc} the average level of human capital of workers in the region and K_{sc} the total amount of physical capital used in the region.

Now, we need to just solve the equation for average labor productivity $\frac{Q_{sc}}{N_{sc}}$:

$$
\begin{aligned}
Q_{sc} &= A_{sc}\Omega_{sc}\left(\left(\frac{N_{sc}H_{sc}}{A_{sc}}\right)^{\beta}\left(\frac{K_{sc}}{A_{sc}}\right)^{1-\beta}\right)^{\alpha}\left(\frac{Q_{sc}}{A_{sc}}\right)^{\frac{\lambda-1}{\lambda}} && |()^{\lambda} \\
Q_{sc}^{\lambda} &= A_{sc}^{\lambda}\Omega_{sc}^{\lambda}\left(\left(\frac{N_{sc}H_{sc}}{A_{sc}}\right)^{\beta}\left(\frac{K_{sc}}{A_{sc}}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{Q_{sc}}{A_{sc}}\right)^{\lambda-1} && |\cdot Q_{sc}^{-\lambda+1} \\
Q_{sc} &= \Omega_{sc}^{\lambda}\Big(N_{sc}H_{sc}\Big)^{\alpha\beta\lambda}K_{sc}^{(1-\beta)\alpha\lambda}A_{sc}^{\lambda}\frac{1}{A^{\alpha\lambda}}\frac{1}{A^{\lambda-1}} && |: N_{sc} \\
\frac{Q_{sc}}{N_{sc}} &= \Omega_{sc}^{\lambda}N_{sc}^{\alpha\beta\lambda-1}(H_{sc}^{\beta}K_{sc}^{1-\beta})^{\alpha\lambda}(\frac{1}{A_{sc}})^{\alpha\lambda-1} && (2.5)
\end{aligned}
$$

By noting that $N_{sc}^{\alpha\beta\lambda-1} = N_{sc}^{\alpha\beta\lambda-1+\alpha\lambda-\alpha\lambda} = N_{sc}^{(\beta-1)\alpha\lambda}N_{sc}^{\alpha\lambda-1} = \frac{N_{sc}^{\alpha\lambda-1}}{N_{sc}^{(1-\beta)\alpha\lambda}}$ we can write the equation (2.5) in the following notation:

$$
\frac{Q_{sc}}{N_{sc}} = \Omega_{sc}^{\lambda}\left(H_{sc}^{\beta}\left(\frac{K_{sc}}{N_{sc}}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{N_{sc}}{A_{sc}}\right)^{\alpha\lambda-1}. \tag{2.6}
$$

It is necessary to simplify the obtained specification of productivity as we don't want to incorporate the K_{sc} variable in the estimating equation. This can be done by using the capital-demand function derived from the equation (2.3):

$$
K_{sc} = \frac{\alpha(1-\beta)}{r_c}Q_{sc}. \tag{2.7}
$$

We get the average labor productivity simply by substituting (2.7) into (2.6) and reorganizing the terms. First, we substitute the $K_{sc} = \frac{\alpha(1-\beta)}{r_c}Q_{sc}$ into the equation:

$$
\frac{Q_{sc}}{N_{sc}} = \Omega_{sc}^{\lambda}\left(H_{sc}^{\beta}\left(\frac{\frac{\alpha(1-\beta)}{r_c}Q_{sc}}{N_{sc}}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{N_{sc}}{A_{sc}}\right)^{\alpha\lambda-1}.
$$

Then, we divide by $Q^{\alpha\lambda(1-\beta)}$ and after that by $N^{-\alpha\lambda(1-\beta)}$

$$
\begin{aligned}
\frac{Q_{sc}^{1-\alpha\lambda(1-\beta)}}{N_{sc}} &= \Omega_{sc}^{\lambda}\left(H_{sc}^{\beta}\left(\frac{\frac{\alpha(1-\beta)}{r_c}}{N_{sc}}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{N_{sc}}{A_{sc}}\right)^{\alpha\lambda-1} \\
\frac{Q_{sc}^{1-\alpha\lambda(1-\beta)}}{N_{sc}^{1-\alpha\lambda(1-\beta)}} &= \Omega_{sc}^{\lambda}\left(H_{sc}^{\beta}\left(\frac{\alpha(1-\beta)}{r_c}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{N_{sc}}{A_{sc}}\right)^{\alpha\lambda-1}.
\end{aligned}
$$

Now we can establish a joint power in the left-hand side of the equation and raise the whole equation to the power of its inverse $\frac{1}{1-\alpha\lambda(1-\beta)}$:

$$
\begin{aligned}
\left(\frac{Q_{sc}}{N_{sc}}\right)^{1-\alpha\lambda(1-\beta)} &= \Omega_{sc}^{\lambda}\left(H_{sc}^{\beta}\left(\frac{\alpha(1-\beta)}{r_c}\right)^{1-\beta}\right)^{\alpha\lambda}\left(\frac{N_{sc}}{A_{sc}}\right)^{\alpha\lambda-1} \\
\frac{Q_{sc}}{N_{sc}} &= \Omega_{sc}^{\frac{\lambda}{1-\alpha\lambda(1-\beta)}}\left(H_{sc}^{\beta}\left(\frac{\alpha(1-\beta)}{r_c}\right)^{1-\beta}\right)^{\frac{\alpha\lambda}{1-\alpha\lambda(1-\beta)}}\left(\frac{N_{sc}}{A_{sc}}\right)^{\frac{\alpha\lambda-1}{1-\alpha\lambda(1-\beta)}}.
\end{aligned}
$$

Finally, we will reorganize the terms:

$$
\frac{Q_{sc}}{N_{sc}} = \Omega_{sc}^{\frac{\lambda}{1-\alpha\lambda(1-\beta)}} H_{sc}^{\frac{\beta\alpha\lambda}{1-\alpha\lambda(1-\beta)}}\left(\frac{\alpha(1-\beta)}{r_c}\right)^{\frac{\alpha\lambda(1-\beta)}{1-\alpha\lambda(1-\beta)}}\left(\frac{N_{sc}}{A_{sc}}\right)^{\frac{\alpha\lambda-1}{1-\alpha\lambda(1-\beta)}}. \tag{2.8}
$$

To make this equation easier to handle, let's simplify it by defining some auxiliary variables. Let:

$$
\begin{aligned}
\Lambda_{sc} &= \left(\frac{\alpha(1-\beta)}{r_c}\right)^{\frac{\alpha\lambda(1-\beta)}{1-\alpha\lambda(1-\beta)}} \\
\omega &= \frac{\lambda}{1-\alpha\lambda(1-\beta)} \\
\theta &= \frac{\alpha\lambda-1}{1-\alpha\lambda(1-\beta)}.
\end{aligned}
$$

Let's also assume that the distribution parameter β is equal to one in the case of human capital. We base this assumption on the notion that education of workers is equally distributed in a region. Using the assumption we get the following form: $H_{sc}^{\frac{\beta\alpha\lambda}{1-\alpha\lambda(1-\beta)}} = H_{sc}^{\frac{\alpha\lambda}{1-\alpha\lambda(1-\beta)}}$.

Now we can write the above equation (2.8) in a cleaner form:

$$\frac{Q_{sc}}{N_{sc}} = \Lambda_{sc}\Omega_{sc}^{\omega}H_{sc}^{\frac{\alpha\lambda}{1-\alpha\lambda(1-\beta)}}\left(\frac{N_{sc}}{A_{sc}}\right)^{\theta}. \tag{2.9}$$

Finally, by performing a sophisticated trick to the exponent of H_{sc}, we'll get it in a slightly different form:

$$\begin{aligned} &H_{sc}^{\frac{\alpha\lambda}{1-\alpha\lambda(1-\beta)}} && \mid \cdot\, H^{1}H^{-1} \\ &H_{sc}^{\alpha\lambda-1+1-1+\alpha\lambda(1-\beta)} && \\ &H_{sc}H_{sc}^{\frac{\alpha\lambda-1}{1-\alpha\lambda(1-\beta)}}. && \end{aligned}$$

Now we can add the newly manipulated H_{sc} term back to the equation (2.9) and see that it forms the equation that will be estimated:

$$\frac{Q_{sc}}{N_{sc}} = \Lambda_{sc}\Omega_{sc}^{\omega}H_{sc}\left(\frac{N_{sc}H_{sc}}{A_{sc}}\right)^{\theta} \tag{2.10}$$

This is the equation for the dependent variable and will be the base for the final estimating equation. In the equation, Λ depends on the rental price of capital in the country and ω is a constant. The most important variable here is the θ that describes the effect of density of employment and human capital in a region on the region's productivity. This effect is the agglomeration effect that is the core of the whole study and the estimation of its magnitude answers directly to the main research question.

To make the average labor productivity function appear in a form that is possible to be used as an estimating equation, logarithms will be taken of both sides of the function and the terms will be reorganized. This allows for a clearer representation:

$$log\frac{Q_{sc}}{N_{sc}} = log\Lambda_{sc} + \theta(log\frac{N_{sc}}{A_{sc}}) + (\theta+1)logH_{sc} + \omega log\Omega_{sc}. \tag{2.11}$$

We don't want to make the model too complex due to the relatively low number of observations. Therefore, two further assumptions will be made: the differences in the rental price of capital Λ can be assumed to be depicted by the spatial fixed effects in the country level and differences in total factor productivity Ω can be assumed to be captured by the regression error term. We also need to use a proxy for human capital as it is really not a quantifiable measure on its own. This will be done using the level of education of the employed population. Applying these changes, the estimating equation takes the form:

$$log\frac{Q_{sc}}{N_{sc}} = \theta(log\frac{N_{sc}}{A_{sc}}) + \sum_{e=1}^{E_c}\delta_e F_{esc} + \text{Country dummies} + u_{sc}. \tag{2.12}$$

As can be seen, we now have an equation that regresses *log* employment density (number of employed people/square kilometer), education level of workers (% of employed population with each education level) and country dummies (1 or 0) to *log* labor productivity (gross value added in Euro/number of employed people). The first estimated coefficient of the first regressor θ quantifies the agglomeration effect, the second coefficient δ_e on the right-hand side estimates the effect of given schooling level and country dummies represent the country fixed effects. We sum the education ratios from the lowest education level e to the highest possible education level E_c in a country c. Unlike in the Ciccone (2002) paper, regional dummies are not used in this study. This is due to the far lower number of total NUTS 3 regions that does not allow to effectively capture differences related to their inclusion in different NUTS 1 and NUTS 2 regions.

It is important to understand the interpretation for the education term as it serves as the main control variable: the δ_e serves as the estimated coefficient for each schooling level and the F_{esc} parameter is the rate of given education level in a region. By noting that we have a sum over different education levels in front of the term, we can see that in this particular case we end up having three education terms, one for each education level. Here, the subscript e means that the estimated coefficient is defined separately for each education level e and the subscript esc means each education level's ratio in each region *sc*.

As the education variable is chosen based on the theory but we do not know for certain how schooling affects productivity, a total of seven separate regressions are ran. This is to empirically study which education levels are most connected with productivity. These regressions are combinations with different included education variables.

2.5.2 The agglomeration effect

As the main interest in the analysis is the estimated value of θ parameter, it is fundamental to discuss the theoretical definition of the agglomeration effect parameter in a little more detail. The θ is defined as follows:

$$\theta = \frac{\alpha\lambda - 1}{1 - \alpha\lambda(1-\beta)}, \tag{2.13}$$

where α determines the returns to capital and labor, λ the externalities from the density of production and β distributional characteristics.

To understand the mechanics of the θ parameter, an illustrative way is to use a few example cases. First, consider the case in which the θ would be zero i.e. there wouldn't be any externalities, no positive nor negative, from the agglomeration of people to the productivity. This would hold if the numerator of the above equation would equal to zero ($\alpha\lambda - 1 = 0$). One combination fulfilling this criterion would be $\alpha = 1$ and $\lambda = 1$ that would mean there would be constant returns to capital and labor and no externalities from the density of production in the region. The same would be the case also when $\alpha < 1$ and $\lambda > 1$ in the way they would balance each other to make $\alpha\lambda$ equal one.

Now, let's consider the case where θ would be positive and the increased density of population would have a positive effect on regional average productivity. This would be true when $\alpha\lambda - 1 > 0$ and so $\alpha\lambda > 1$. From the equation (2.13) we see that the numerator being over one the overall θ value will be greater the greater is the difference in $1-\beta$. β being the distribution parameter, this can be interpreted that smaller the β the larger the agglomeration effect. As β is defined between 0 and 1, its value being 1 would mean completely equal distribution across the region and the agglomeration effect would be minimal. According to the theory developed by Ciccone (2002) the physical capital moves to more productive regions and therefore reinforces the effect more when there are large disparities between regions.

In the last possible case the agglomeration effect θ would be negative. This kind of effect would arise in a situation in which negative congestion effects would dominate positive externalities so much that $\alpha\lambda < 1$. This kind of radical situation would most probably take place in the event of extreme urbanization gone too far where the urban area would have grown too fast and uncontrollably. In this kind of event the increase in agglomeration would lead to negative effect in productivity.

2.6 Empirical results in the literature

The results obtained in the Ciccone and Hall (1996) paper using nonlinear least squares estimator for the agglomeration effect yield 5.2 % increase in the labor productivity for the US sample with a 0.008 standard error. The IV estimate for the agglomeration effect yields 6 % increase in the labor productivity using all four aforementioned instruments with a standard error of 0.01.

In the nonlinear least squares regression, the goodness of fit $R^2 = 0.551$ and in the IV $R^2 = 0.536$. As both of these can be intepreted as moderate, it is important to keep in mind that the variance in the incidators is fairly high and the R^2 itself tells very little of the validity of the estimations.

In the Ciccone (2002) paper on European countries estimates for the agglomeration effect (θ) are somewhat lower. The OLS regressions yield values of 5.058 % (country dummies), 5.07 % (NUTS 1 region dummies) and 4.97 % (NUTS 2 region dummies). Correspondingly, the IV estimates yield 4.55 % (country dummies), 4.445 % (NUTS 1 region dummies) and 4.444 % (NUTS 2 region dummies)[1]. These all have their White adjusted standard errors in the range of 0.417 % - 0.592 % and no evidence of different values for agglomeration effect between countries is found.

In the light of these results, it can be expected that the estimates for the Nordic

[1]NUTS 1 regions e.g. in Finland comprise of two regions: Mainland Finland and Åland islands. Correspondingly, NUTS 2 regions in Finland comprise of five regions: East, South, West and North Finland and the Åland islands.

countries would produce similar or slightly lower values. It has to be noted also that where the US has 50 states and 3142 counties and Germany has 401 NUTS 3 regions, all Nordic countries only have 72 NUTS 3 regions in total. This will probably cause some differences between the estimates of these two studies and this research.

3. Estimation

3.1 Methodology

As discussed in the previous chapter, this thesis is based on the model developed by the economists Antonio Ciccone and Robert E. Hall and used in their studies on economies of agglomeration. Most notably, these include a research on the phenomenon in the USA, using county and state level data (Ciccone and Hall 1996) and in the five most populous countries of Europe, using NUTS 3 level data (Ciccone 2002). This study aims at finding out if a similar model specification would provide similar results in the Nordic Countries.

The model formulation starts from the assumption that by increasing the employment density, positive externalities can be experienced in the regional productivity. Put simply, this means that as the number of employed people increases in an area, the productivity of the area increases correspondingly. One way to find out the magnitude of this effect is to study the historical values of economic and socio-economic indicators associated with this phenomenon and draw conclusions from the results.

As the geographical region of interest consists of only four countries and 70 regions in total, the model cannot be too complex. This might be the most restrictive requirement of all since it forces to choose only the most relevant variables and argument this decision thoroughly.

3.1.1 Estimating equation

The estimating equation that will be used in this study was formulated in the chapter 2.5.1. The result we got is the equation 2.12 that can best be described as an equation for a linear regression model taking the following form:

$$log\frac{Q_{sc}}{N_{sc}} = \theta(log\frac{N_{sc}}{A_{sc}}) + \sum_{e=1}^{E_c} \delta_e F_{esc} + \text{Country dummies} + u_{sc}.$$

It is important to remember that this equation follows a $log - log$ specification that means the estimates should be interpreted accordingly. The parameter of interest here is the θ as it describes the magnitude of the agglomeration effect in the Nordic countries. Before estimating the models it is, however, good to explore the data in a little more detail to get the understanding of the region we are dealing with.

3.2 Data

In Econometrics, data forms the base of every analysis and therefore special attention must be pointed to the quality and appropriateness of it. As this research concentrates on studying agglomeration effects based on regional indicators, it requires data collected on NUTS 3 level, put more simple, in a level of finest detail collected in Europe.

In this study, a cross-sectional approach has been chosen based on the observed low variance between the years of the included indicators. This is mainly due to the fact that regional demographics change slowly. It needs to be stated at this point that one of the selection criteria for the particular geographic area was the good availability of data. Nordic countries are on the top of the list when it comes to the quality and availability of data. In addition to the data found at the Eurostat[1] database, each country publishes fairly clean, open data in their local statistics institutes' websites[2]. As dictated by the previous section's estimating equation, data needed are production, employment, land area and education level of workers, all in NUTS 3 level.

For production, Gross Value Added (GVA) is used as a proxy for the production as is proposed by the OECD (2019). This data is available in the Eurostat (2018c)

[1] *The Statistical Office of the European Communities* is a statistical office with a responsibility to gather and govern data from all European Union member states.

[2] These include *Statistics Finland* in Finland, *Statistics Sweden* in Sweden, *Statistics Norway* in Norway and *Statistics Denmark* in Denmark

database and based on the completeness of the data series for the countries in question, year 2015 is chosen being the most recent year all data are available.

For employment, meaning the number of employed people in a given region, data is available also in the Eurostat (2018b) database in the NUTS 3 granularity. The land area of regions is gathered from the same place (Eurostat 2018a) as it can be assumed to be the most precise and reliable source.

In the case of education levels, the story goes a little differently. As all countries have slightly different schooling systems, their comparison is not straightforward. As an example of this, Denmark seem to have quite an impressive number of different options for secondary schooling that are not at all unambiguous (The Child and Youth Administration, City of Copenhagen 2019). However, a final decision to restrict the number of education levels to three makes this easier as it allows to aggregate all these different schooling options into one category.

Finland has without a doubt the best quality of data on education of population: as an example, higher education is parsed in groups of lower, higher, doctoral and post-doctoral levels and the values add up nicely (Statistics Finland 2018). This is not to say that the Swedish (Statistics Sweden 2018), Norwegian (Statistics Norway 2018) or Danish (Statistics Denmark 2018) statistics institutes would not produce the data of similar quality but to point out the interesting fact that maybe the pervasive pride in the high standard of the education system is also visible in the data gathering process.

Final percentages of education levels of employed people have been calculated based on these data on the number of educated workers in a given region and divided by the total number of region's workers. This has enabled us to form a harmonized and intercomparable measure for education across all regions.

3.2.1 Variance of explanatory variables

The function of explanatory variables in the regression is to provide information on the phenomenon studied. The estimation is made based on these data and therefore it is fundamental that the variance the explanatory variables exhibit is not too small. This is the reason we need to inspect the distribution of these variables.

This inspection is most easily done by plotting the histograms of these variables. An ideal situation would be that all of the variables (excluding the dummies) would have a wide spread of values that would provide a lot of information for the basis of regression.

As seen in the figure 3.1, the first variable, *log* employment density, gets values that seem to be pretty normally distributed between -0.1868 and 7.9664. This points to a large variance in the employment densities within the dataset which is expected as the sample includes the capital regions of the Nordic countries but also the least populated NUTS 3 region in the whole Europe, Lapland. In addition to that, of the 15 least densely populated NUTS 3 regions seven are included in this sample.

Inspecting the education variables, we can also see a pretty vast range of values. The variable for the ratio of employed population having primary education, gets values between 0.17 and 0.37 showing a 20 percentage point deviation. In the case of the secondary schooling variable, the values range from 0.3 to 0.51 and show a 0.21 percentage point deviation. Lastly, the tertiary schooling variable gets values between 0.2 and 0.48 and experiences a 0.28 percentage point deviation between the most extreme values.

It can be concluded that the underlying data exhibits enough variance to be effectively used for this regression. A table of all values for the economic indicators in the underlying dataset is presented in the appendix.

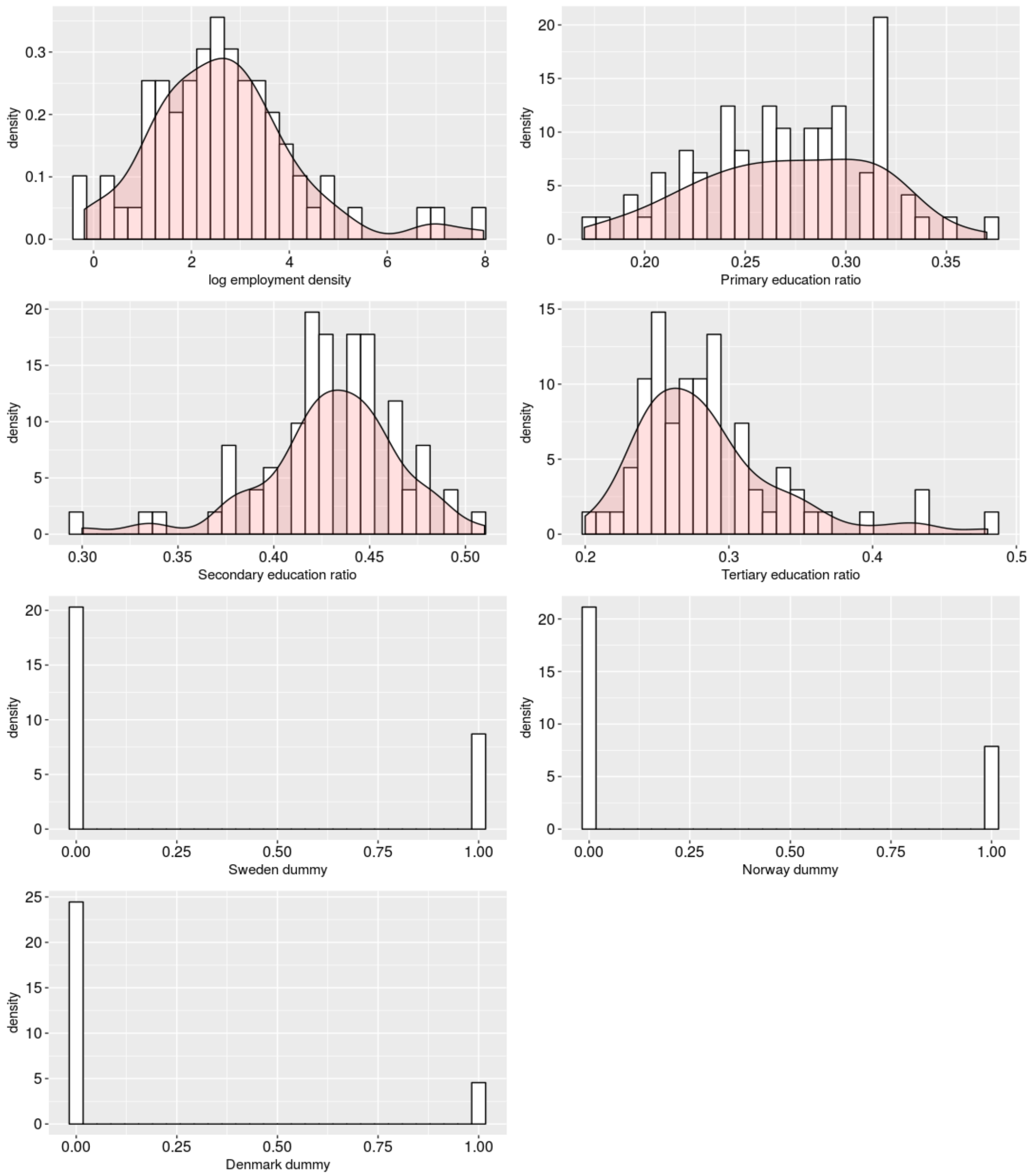

Figure 3.1: Variance of explanatory variables

3.2.2 Descriptive analytics

As the NUTS 3 level regions, counties from now on, might not be clear as a concept for everyone, the emphasis of this section is in map representations. The maps in this section have been made using the open GeoJSON geographical NUTS 3 border data from Datahub (2018) and compiled using the statistical software R. Variable values in the original dataset have been used as a basis for the color codings in many of the maps.

Looking at the map of the Nordic Countries with highlighted county borders[3] in the figure 3.2, an important observation is that the variance in areas of counties is pretty large. Historically, these administrative regions have perhaps been established to allow for roughly the same amount of people to live in one region for easier governance but, as can be seen from the Figure 3.3, the case is not like that anymore.

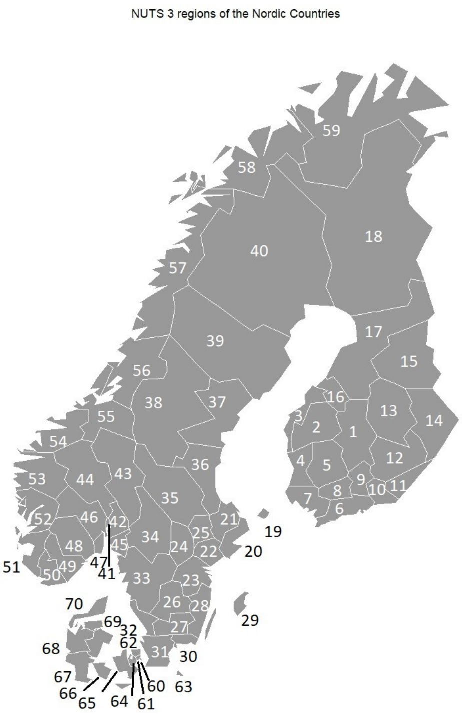

Figure 3.2: NUTS 3 regions of the Nordic countries

The situation in the Nordics is a little different than in Europe which poses a challenge for the empirical discussion of reasons for the prevalent demographics. For example, Finland experienced the most rapid urbanization not before than around the 1960-1970 while in the 1950s most of the population still resided in the countryside. This same development can be observed to have happened in the other Nordic countries also, just a few decades earlier. The population centers mostly to the southern parts of the countries that supports the interpretation of geographically categorizing the Nordic countries to the more peripheral area of Europe.

[3] Indexation of the regions can be found in the appendix

As we use the total number of employed people in a given NUTS 3 region as a variable in our empirical model, it is more sensible to visually inspect the distribution of population in a finer level. The map in the figure 3.3 conveniently shows the distribution of people in the Nordic countries included in the analysis. It can be seen that most of the people indeed live in the south of each country with the exception of Denmark, where people seem to have distributed more equally around the whole country.

Another interesting thing visible in the map is that the Nordic Countries are actually pretty sparsely populated outside the largest cities. If we compare the population density of these countries to Lithuania for example, we see that in Lithuania the population forms a network of densely populated small cities unlike in the Nordics. We can also observe that the densities of Stockholm and Copenhagen clearly exceed Helsinki's and Oslo's densities.

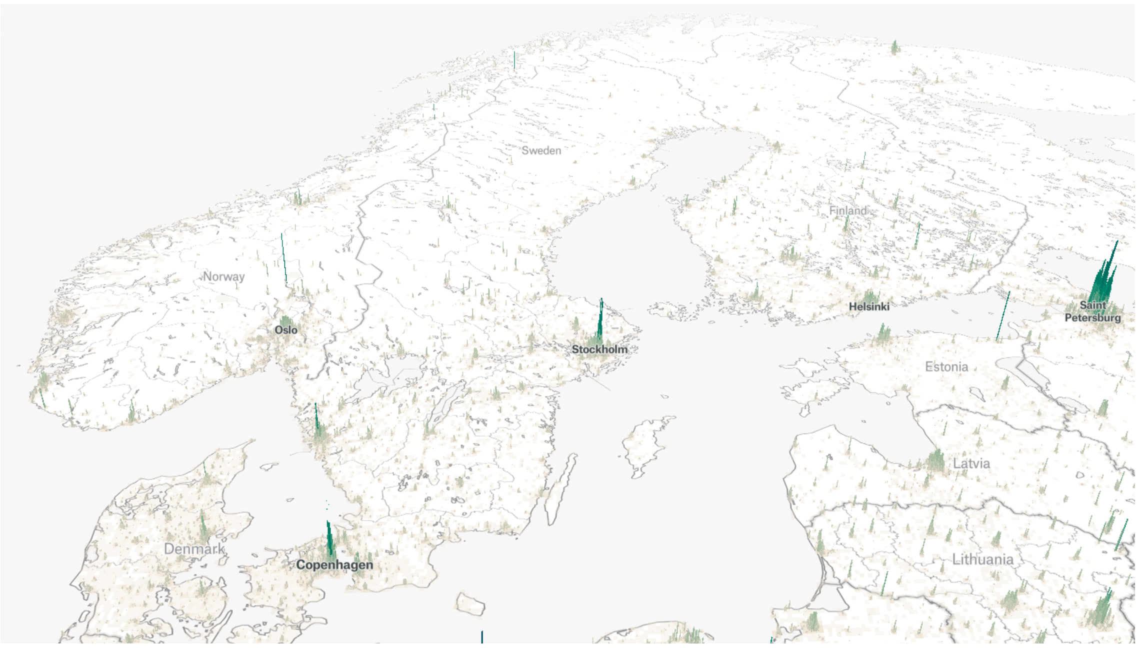

Figure 3.3: Geographical distribution of population across the Nordic Countries (Pudding 2018)

Taking a look at the map showing the distribution of production in Figure 3.4, the main centers of production are easily visible. In Finland, four regions containing the six largest cities consist of Helsinki region, Southwest Finland, Pirkanmaa and Northern Ostrobothnia and the role of these six cities as the driver of economic development is pretty evident.

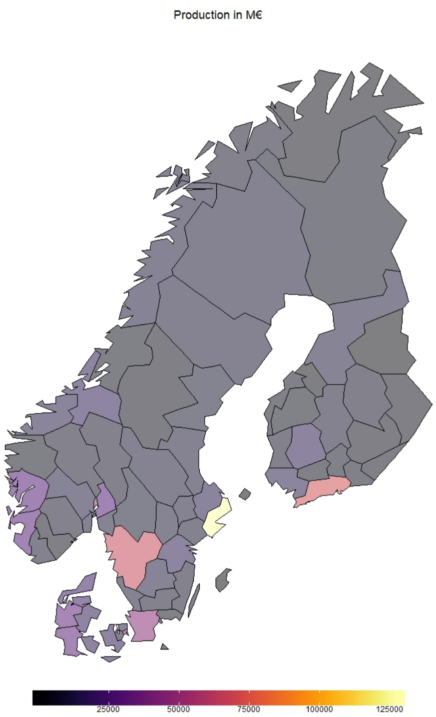

Figure 3.4: Geographical distribution of production across the Nordic Countries

When it comes to Sweden, one important thing to remember is the effect of climate. Roughly third of Sweden's area is located more south than the southernmost point of Finland and the population has understandably accumulated there. This has also had an effect on the distribution of productivity; the most productive regions are located in the southern regions. However, due to Sweden's more vast industry base, there are also productive regions in the northern parts of the country.

Norway, on the other hand is another story. The rise of offshore fishing and oil industries during the past decades have distorted the country's production to more northern regions as production in these fields is heavily tied to location. However, this same trend is not visible in the employment density graph which is evidently due to the massive commuting of the workers that is characteristic for these particular fields.

In the case of Denmark, the discussion of these indicators takes a full turn. Denmark cannot be considered having the same geographical characteristics as the other Nordic Countries as it's area alone is substantially smaller than the others'. The geographical fact that each point in Denmark is not more than 50 kilometers from the seashore makes the shipping cheaper and its land proximity to European markets

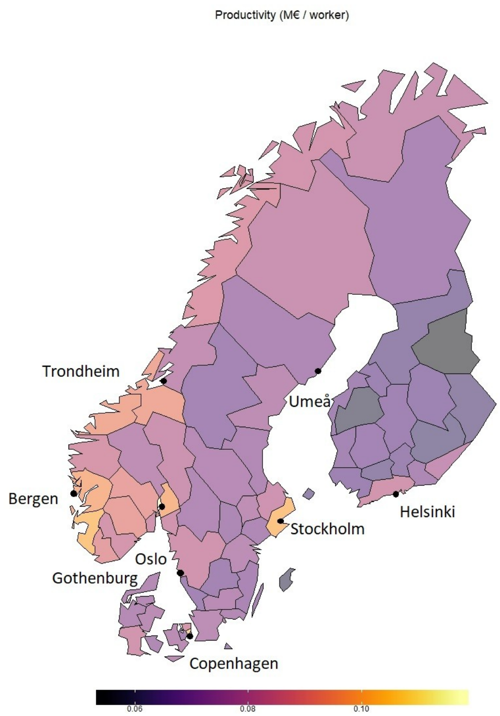

Figure 3.5: Geographical distribution of productivity across the Nordic Countries

benefits the production in all fields.

Being so small in area, Denmark's regional division is denser and people have accumulated smoother across the whole country. This makes it hard to compare population densities with its neighbors and points towards the fact that it should almost be considered to geographically belong to the core of Europe. However, Denmark shares similar history and institutions as other Nordic Countries and is therefore included in this study.

Inspecting the productivity map in Figure 3.5, a concentration of high productivity can be observed in the regions consisting of the capital regions of each country. This hardly strikes as a surprise and is well in line with the theory of economies of agglomeration. There are also other agglomerations of high productivity, such as surroundings of Bergen and Trondheim in Norway and the Gothenburg area in Sweden. These are all areas with access to the sea and are more densely populated as other, more northern, regions. In the north of Sweden, there are also two regions, Jämtland and Västerbotten, that have higher productivity as their neighboring regions. For Jämtland, this might be because of Åre's large ski resort that employs a lot of seasonal workers that generates high levels of production without affecting the employment in the area. The case of Västerbotten, in turn, might result from

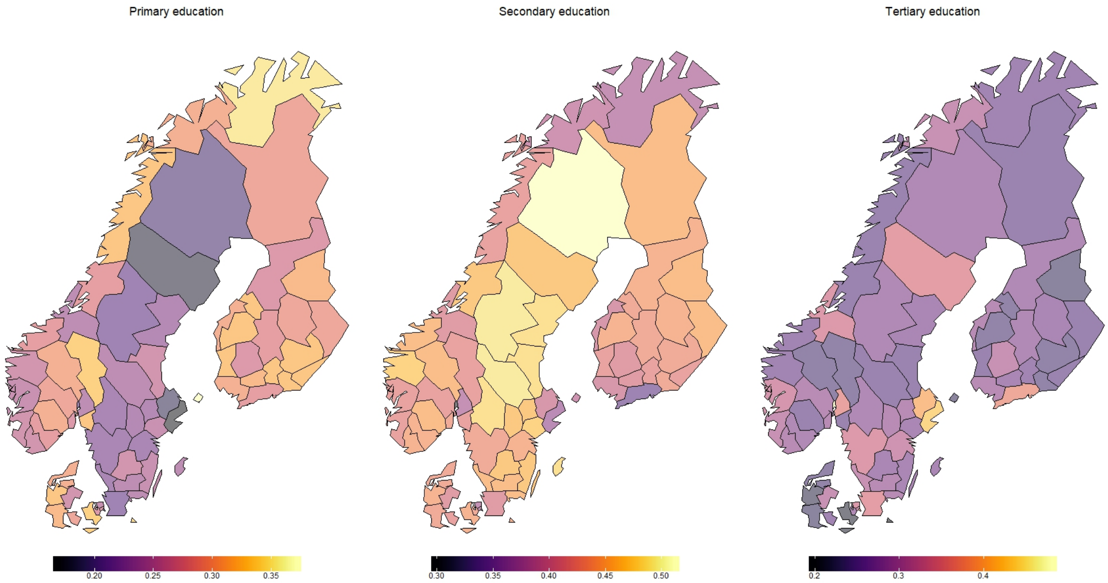

Figure 3.6: Geographical distribution of education across the Nordic Countries

the concentration of technology industries within the Umeå university.

Not to undermine the other indicators, most interesting maps of all are the education maps shown in figure 3.6. Albeit, all the Nordic Countries share an impressive equality of education across all regions, there are some characteristics visible supporting the theory of economies of agglomeration. Highest percentages of the employed population having only a primary education seem to take place in the remotest regions, namely the Northernmost region in Norway and the mountainous region of Hedmark close to the Swedish border in the south.

Taking into account the fact that the lowest unemployment rate in the whole Europe can be found in the island of Åland between Finland and Sweden, its large number of employees having only primary education might first seem a little odd. This might, however, be explained by the significance of fishing and maritime industries as they have also historically been important in the region.

Large percentage of workers with secondary education seem to be especially prevalent in the northern regions of Sweden that can maybe result from the concentration of manufacturing companies in the area. Inspecting the distribution of tertiary education, it is evident that the capital regions pack most of the highest educated people. Västerbotten in the North of Sweden and Sør-Trøndelag in the middle of Norway can be seen as some outliers contrasting the assumed distribution. Once again, cities of Umeå and Trondheim explain this observation pretty well.

4. Results

As our data consists of observations from only the year 2015, we are relying on cross-sectional approach in our regression analysis. We will run a series of OLS regressions incorporating different combinations of variables and choose the most suitable specification based on statistical tests and subject knowledge.

A use of linear regression model is proposed by the existing literature that makes it an intuitive specification to use also in this study. First thing to do when performing a linear OLS estimation is to check if the linearity condition for the OLS estimator holds. We start the model validation from this as it is fundamental to make sure relationships we are studying with linear model actually are linear. This can be done easily by visually inspecting the relationships between each explanatory variable and the dependent variable.

In the figure 4.1 the dependent variable log productivity is plotted against each of the four explanatory variables. The black, hollow dots represent the observations in the data and the regression line represents the trend of the data. From the figure we can easily recognize the relationship between the explanatory variables and the dependent variable to be linear in all cases. There are no signs of quadratic or cubic relationship between the variables that is well in line with our proposed model specification.

In the following section the model results will be gone through and contrasted to the empirical results obtained in the literature for other groups of countries. After that, all fundamental OLS assumptions are addressed and some proof for the validity of chosen model specification will be shown. In the end of this chapter, the results from an alternative regression studying the differences in agglomeration effects between countries will be presented.

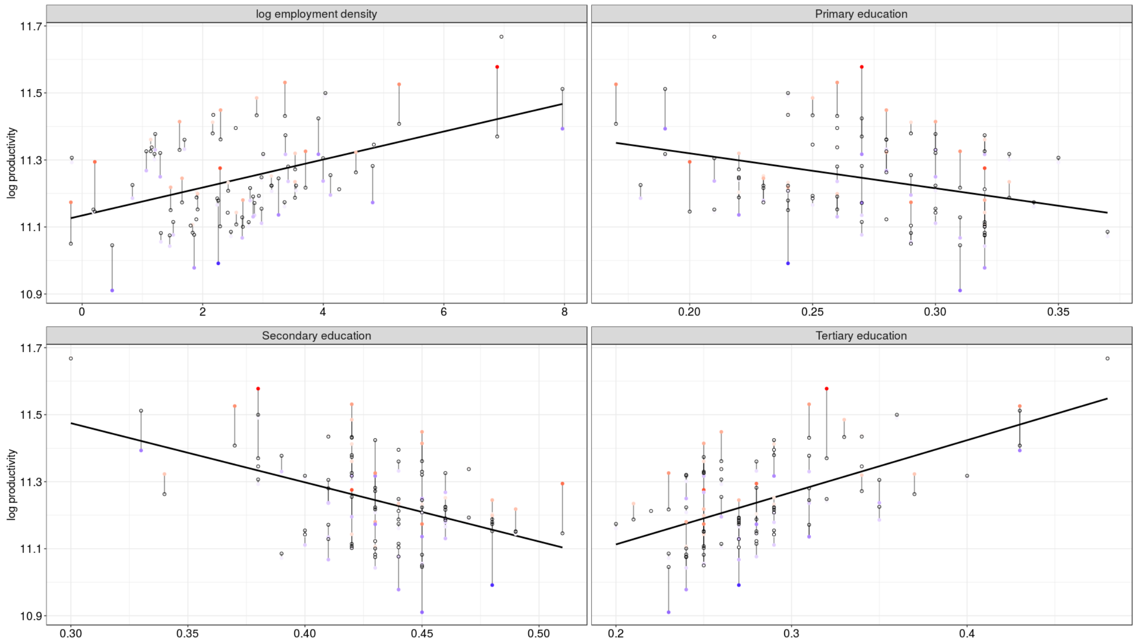

Figure 4.1: Relationships between the dependent variable and the explanatory variables

4.1 Estimation results

This section presents the regression results obtained by running seven different model specifications to study the effect of employment density on regional productivity. Each one of them differ in the included education control variables and yield slightly different results.

The results of seven independent regressions are shown in the following table 4.1. As can be seen in the first row, all models except the model 1 seem to produce very robust results for the θ parameter ranging from 0.021 to 0.029. This can be interpreted as an indication of robustness across all model specifications and it also tells that the inclusion of different levels of schooling in the model do not affect the results too much. Things to note are also the statistical significance of all of these estimates at the 10 % level and fairly small standard errors.

Table 4.1: Regression Results

	Dependent variable:						
	log productivity						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log employment density	0.038^{***}	0.029^{**}	0.022^{**}	0.021^{*}	0.022^{**}	0.022^{*}	0.021^{*}
	(0.008)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
Primary schooling	-1.021^{***}			-0.954^{***}	-0.074		-0.716
	(0.346)			(0.338)	(0.558)		(1.694)
Secondary schooling		-1.032^{**}		-0.922^{**}		-0.005	-0.686
		(0.447)		(0.426)		(0.562)	(1.707)
Tertiary schooling			0.964^{***}		0.918^{**}	0.962^{***}	0.247
			(0.259)		(0.431)	(0.345)	(1.726)
Sweden dummy	0.011	0.133^{***}	0.064^{**}	0.055	0.059	0.064	0.057
	(0.036)	(0.031)	(0.025)	(0.041)	(0.042)	(0.039)	(0.043)
Norway dummy	0.246^{***}	0.270^{***}	0.248^{***}	0.249^{***}	0.247^{***}	0.248^{***}	0.249^{***}
	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)	(0.026)	(0.026)
Denmark dummy	0.051	0.084^{**}	0.101^{***}	0.086^{**}	0.099^{**}	0.101^{***}	0.090^{*}
	(0.034)	(0.039)	(0.036)	(0.037)	(0.040)	(0.037)	(0.047)
Constant	11.343^{***}	11.484^{***}	10.810^{***}	11.747^{***}	10.843^{***}	10.812^{***}	11.508^{***}
	(0.116)	(0.207)	(0.058)	(0.218)	(0.261)	(0.311)	(1.676)
Observations	70	70	70	70	70	70	70
R^2	0.742	0.730	0.759	0.760	0.759	0.759	0.760
Adjusted R^2	0.722	0.709	0.741	0.737	0.737	0.736	0.733
F Statistic	36.841^{***} (df = 5; 64)	34.553^{***} (df = 5; 64)	40.395^{***} (df = 5; 64)	33.248^{***} (df = 6; 63)	33.149^{***} (df = 6; 63)	33.137^{***} (df = 6; 63)	28.058^{***} (df = 7; 62)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Comparing the θ estimates obtained for the Nordic countries with the two empirical papers discussed in the previous chapter, we find out that the estimates for the θ parameter in all model specifications have positive signs. This tells that a positive relationship really exists between the employment density and the productivity and means we are not dealing with congestion effects in the Nordics nor experience any other large negative effects. This finding conveniently confirms our first hypothesis for the study: there actually exists a positive agglomeration effect in the Nordics.

When it comes to the magnitude of the effect, it can be interpreted way smaller than the effect in the US or in the largest European countries. As our model uses a $log - log$ specification, one percent increase in employment density will induce a 0.021 - 0.038 % increase in the productivity and, correspondigly, doubling the employment density (100 % positive change) will raise the productivity by 2.1 - 3.8 %. Remembering the θ value for the US being 5.2 - 6 % and for the largest European countries 4.4 - 5.07 %, Nordic values are significantly lower.

The control variables are also interesting to interprete: they seem to also get fairly robust values across all specifications but experience high standard errors in some model specifications that affect the statistical significance. However, the signs seem to get uniformly negative values for the primary and secondary schooling and positive values for the tertiary schooling variable. The implication of this is that, according to the models, high levels of tertiary schooling have a positive effect on productivity but as the ratio of primary and secondary education increases, the productivity decreases. This is well in line with the literature and the whole model specification as the accumulation of human capital is expected to drive the productivity in the region.

Inspecting the estimated values for the country dummy variables, one thing pops up: the country dummy for Norway seems to be statistically significant at the 1 % level across all model specifications. Given that the value is positive, it certainly seems that there is something fundamentally different in Norway that affects the productivity level. This might result from the economic characteristics addressed in the previous chapter e.g. abnormally profitable petroleum and fishing industries within the Nordic countries.

Other statistics such as the goodness of fit R^2 and the adjusted R^2 report a fairly good fit between the data and the model specifications showing values between 0.709 and 0.760. The F statistic that measures the joint significance of parameters gets also high values for all models suggesting that the parameters are jointly statistically significant at the 1 % level in all specifications.

4.2 OLS assumptions

For these results, we used a linear OLS model. As we now have multiple good model candidates for the final results it is important to check whether these models actually fulfill all the OLS model assumptions that enable us to trust the obtained estimates. These assumptions are being tested and evaluated in the following four subsections.

4.2.1 Exogeneity of explanatory variables

As maybe the most important thing determining the model specification between OLS and IV regressions is the potential endogeneity of our explanatory variables, it is important to test this right in the beginning. Endogeneity arises when the explanatory variable is correlated with the error term and is therefore affected by some external characteristics that are not being taken into account by the model. If there is enough evidence to suspect this kind of a situation, a common way to deal with it is to use an instrumental variable that will capture these characteristics and thus make the explanatory variable independent.

There is a reason for testing this possibility: we don't want to use IV regression just for the sake of it as IV estimators often have larger standard errors compared to the OLS estimators. Therefore, if we find the explanatory variable to be exogenous, the OLS estimator will be the best linear unbiased estimator and also our estimator of choice.

As proposed in the literature by Ciccone (2002), a good instrument for the explanatory variable *log* employment density would be the total land area of NUTS 3 regions. This has proven to be effective in the study of European countries and we will also rely on it in this study. It is not possible to be statistically certain that the chosen instrument is relevant and in the literature this part is often assessed by discussing the historical or other characteristics of the phenomenon of interest. In this study, we rely on Ciccone's reasoning of historical goal to equalize the population among the regions for administrative purposes.

Now, we will test whether the explanatory variable is exogenous or not. The result of this Hausman test will determine if we can use the OLS as our model specification or if we need to resort to the IV approach. The null hypothesis of the Hausman test is that both OLS and IV regressions give consistent estimates and in the case it will be rejected, we will have to resort to IV estimators.

Technically, we will first regress our main explanatory variable to the other variables and the instrumental variable and perform this for all models. After that we will run our original regression between *log* productivity and all original explanatory variables added with the residuals from the previous regression. Then, by performing the Wald test between the original regression with and without the first stage residuals we can see if the null hypothesis of these residuals being irrelevant can be rejected.

The below table 4.2 of the Hausman test results points to the rejection of the null hypothesis. According to the results, the dependent variable experiences endogeneity in all model specifications at the 5 % level. Based on this, IV regression would be the correct model specification to continue.

Table 4.2: Hausman test for dependent variable exogeneity

Model	Test statistic (F)	p-value
1	9.6038	0.002901 **
2	12.969	0.0006247 ***
3	7.9293	0.006486 **
4	7.7949	0.006957 **
5	7.8028	0.00693 **
6	7.8718	0.0067 **
7	7.6529	0.007491 **

When the IV regressions were performed, they yielded very varying estimates. As an example, the estimates showed both negative and positive values for the explanatory variable and also for all control variables. These results also exhibited abnormally large standard errors, in some cases even larger than the estimate itself. Based on these findings, there is no other choice than to consider the chosen instrument to be invalid in this particular geographical region.

Another thing supporting this interpretation would be the robustness of the results in the OLS regressions across all model specifications. This tells that it is not probable that there would be something fundamentally wrong with the model specification.

4.2.2 Homoskedasticity

After validating the exogeneity condition for the model, we need to make sure there is no heteroskedasticity present in the model residuals that means the error terms would exhibit a constant variance over time i.e. would be homoskedastic. This is most easily done by plotting the model residuals against the fitted values predicted by the model. In the figure 4.2 plotted model residuals seem to have a fairly constant variance with a zero-mean, a result that allows us to consider this OLS assumption satisfied. This result implies the presence of homoskedasticity and lets us to continue without resorting to means of assessing heteroskedasticity e.g. use White standard errors.

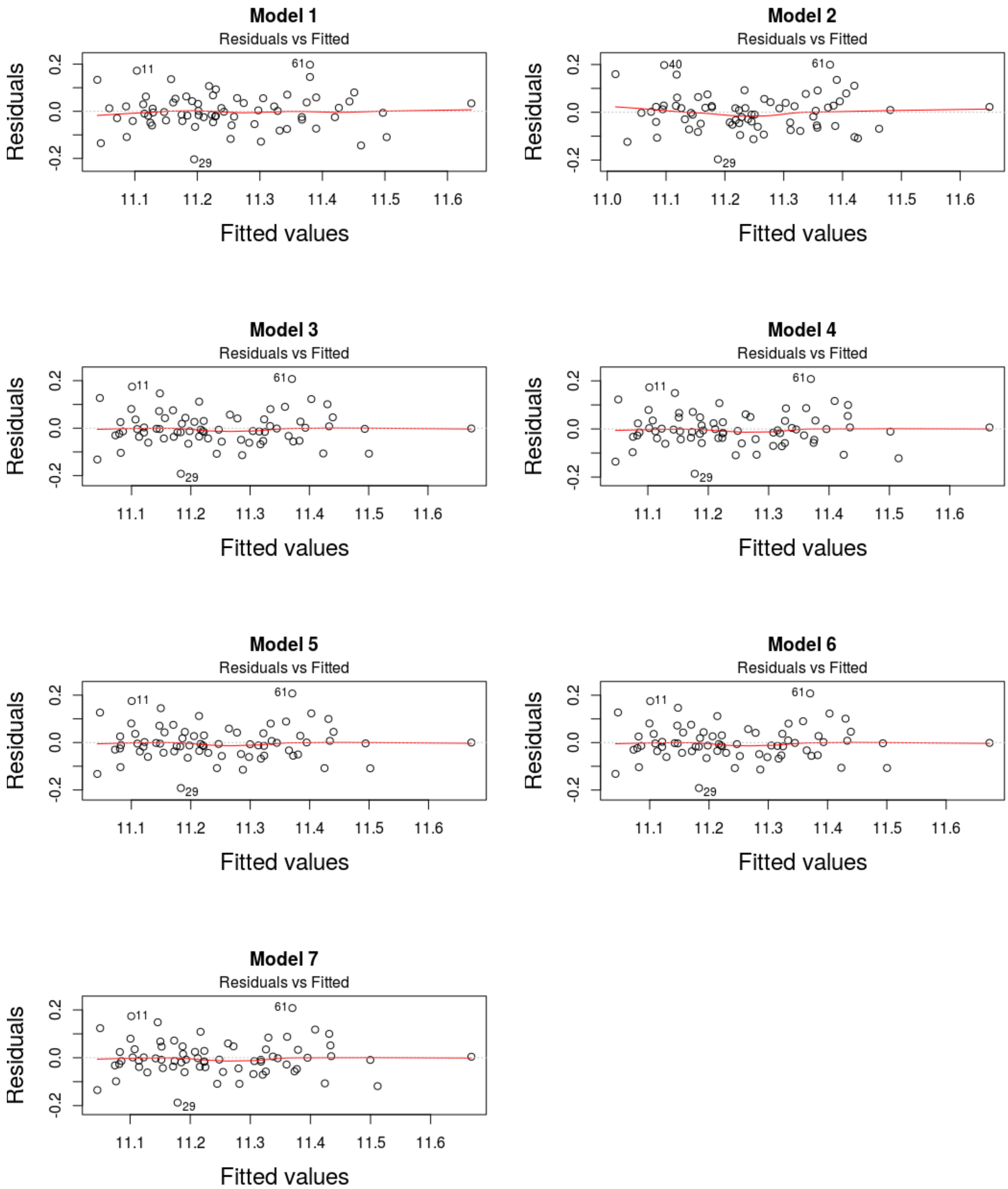

Figure 4.2: Model residuals against the fitted values

By plotting the histograms of the standardized residuals we can also inspect the normality of residuals. The figure 4.3 shows the distribution of different values obtained by the model residuals. It is easily visible that the residuals in each of the models roughly follow the normal distribution.

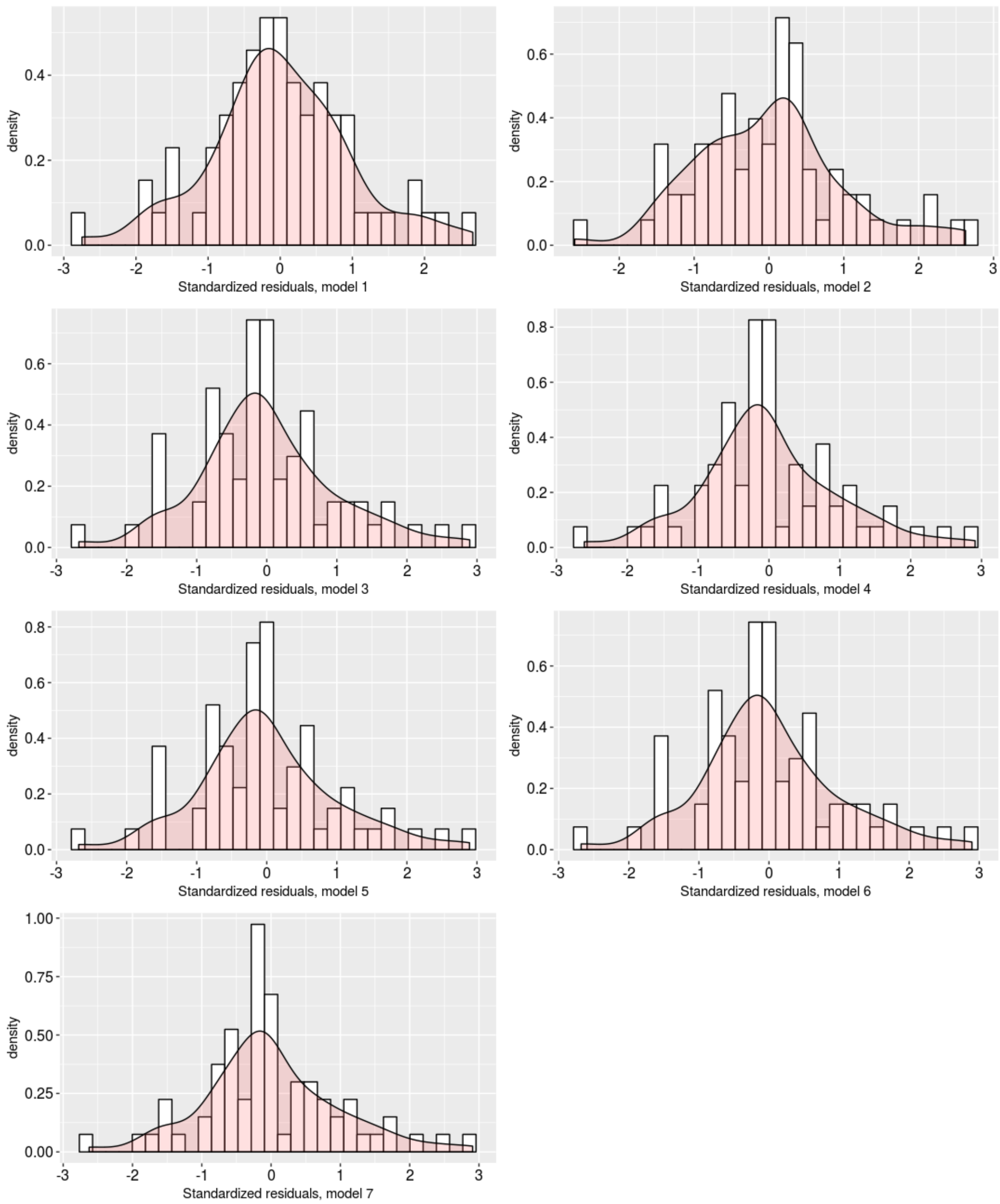

Figure 4.3: Normality of residuals

Based on these results we can finish the discussion on homoskedasticity by concluding the error terms to be homoskedastic in all models. Next, we move on to examine the possible multicollinearity between the explanatory variables.

4.2.3 Multicollinearity

We will now examine the presence of multicollinearity that arises if there are linear relationships between the independent variables of the model. Multicollinearity is an issue since the correlation between explanatory variables violates one of the fundamental properties of OLS, the independence of explanatory variables. Let's first test this by calculating the Variance inflation factors (VIFs). The VIF values are shown in the table 4.3.

Table 4.3: VIF test results

Model	log(employment/area)	primary	secondary	tertiary	Swe dummy	Nor dummy	Den dummy
1	1.829549	2.740772	3.259999	1.598294	1.831409		
2	3.343538	3.034992	2.325229	1.464034	2.249451		
3	3.011157	2.291251	1.619975	1.534007	2.208723		
4	3.600843	2.763871	3.060570	4.341009	1.604760	2.250187	
5	3.386657	7.508599	6.277095	4.602212	1.600111	2.664046	
6	3.559496	5.316906	4.013971	3.935460	1.620536	2.314692	
7	3.626379	68.235431	48.318121	99.098334	4.707154	1.624315	3.543238

A critical value for the presence of high multicollinearity is considered to be VIF > 5. These results show that models 5 , 6 and 7 have VIF levels over 5 and therefore exhibit high multicollinearity. This is a clear violation of the OLS assumptions so the usage of these models will not be an option without some serious respecifications.

Multicollinearity can be further studied by forming a correlation matrix (table 4.4) of correlations between the explanatory variables:

Table 4.4: Correlation table of explanatory variables

	log.GVA.employment.	log(employment/area)	primary	secondary	tertiary	Swe dummy	Nor dummy	Den dummy
log.GVA.employment.	1	0.442	-0.316	-0.447	0.558	-0.119	0.607	0.097
log(employment/area)		1	-0.257	-0.637	0.564	-0.081	-0.14	0.541
primary			1	-0.078	-0.727	-0.698	0.138	0.132
secondary				1	-0.609	0.473	-0.17	-0.211
tertiary					1	0.215	0.074	-0.074
dummy_swe						1	-0.4	-0.283
dummy_nor							1	-0.264
dummy_den								1

This table tells the same story: primary and tertiary education variables and secondary and tertiary education variables are highly correlated. This points to the fact that these variables cannot be used together in the same model without violating the OLS assumptions.

4.2.4 Functional form of the model specification

To back up our initial visual inspection in the figure 4.1, we will test the model specification using the Ramsey's Regression Equation Specification Error Test (RESET) against the quadratic and cubic model specifications. In this test, the null hypothesis is that we have specified the model correctly (in this case, to be linear) and the rejection of it would point to the need to respecify the model in higher order.

The table 4.5 introduces the results of the RESET tests performed for all seven different model specifications. It is clear that none of the results of the RESET test points to the rejection of the current specification as the p-values clearly exceed even the 10 % threshold. Based on these two validations, we can continue with the linear model specification. The actions required in the light of the other test results will be discussed in the following section.

Table 4.5: Ramsey's RESET test results

Model	RESET test statistic	df 1	df 2	p-value
1	0.63316	10	54	0.7789
2	0.88803	10	54	0.55
3	0.29807	10	54	0.9787
4	0.57397	12	51	0.8527
5	0.28807	12	51	0.9888
6	0.61444	12	51	0.82
7	0.55517	14	48	0.8856

4.3 Choosing the valid model specifications

In the light of previous tests' results, the presence of multicollinearity have proven to pose a constraint for the usage of some of the more extensive models. In an economic sense this was expected as all controls measure essentially the same thing: education of population. In this particular case, inclusion of three different education levels is not necessary as education levels are fairly uniform across the regions; it would be another story in a sample of countries where disparities in education would be essentially larger.

Based on the results in the VIF table (table 4.3), there seems to be evidence towards the exclusion of models 5, 6 and 7. The VIF values exceed the threshold for high multicollinearity (VIF>5) by large and models can be interpreted to produce biased estimates. Further evidence pointing towards the exclusion of these models is the statistical insignificance of many of their parameters, such as primary schooling in model 5, secondary schooling in model 6 and almost all parameters in model 7.

When it comes to economic reasoning, controlling only population's primary schooling does not seem sensible as the theory behind the model assumes the productivity to be driven by the accumulation of human capital in a region. Based on this, we can also exclude the first model.

We are now left with models 2, 3 and 4. All their estimates are within the range of 2.1 and 2.9 % that seems sensible contrasting to the previous studies performed on the topic in the USA and in other European countries. In the models 2 and 3, all estimates seem to be statistically significant at the 5 % level and also the R^2 values seem quite high. To further argue the correctness of the models, all models exhibit statistically significant F statistics that imply the variables to be jointly significant.

4.4 The magnitude of agglomeration effect between countries

So far, we have concentrated on studying the agglomeration effect across all Nordic countries that has yielded results ranging between 2.1 and 2.9 %. In this section, we will form a regression with an aim of finding out if this effect gets significantly different values between individual countries.

The model specification used so far bases on the assumption that Nordic countries are quite a homogeneous group and it would be possible to find a value that holds for all of them. This has been the main objective of this study and have proven to be possible in the previous section. However, every good study requires a sensitivity analysis and the following will work as the one for this.

The results of the regressions with country level agglomeration effects included are presented in the table 4.6. These have been done by adding three completely new regressors in the three model specifications (2, 3 and 4) that have proven to satisfy all of the OLS assumptions. These three new regressors are essentially the same as the explanatory variable (*log* employment density) but are supplemented by the country dummies as their coefficients. This way each of these estimators gets values only when each country's dummy variable gets a value one and thus measures the difference in agglomeration effects between that particular country and Finland. Due to this, we cannot interpret the coefficients as the effect itself, only the difference in effects compared to Finland.

Table 4.6: Alternative Regression Results

	Dependent variable:		
	log productivity		
	(1)	(2)	(3)
log employment density	0.024	0.029	0.029
	(0.021)	(0.019)	(0.021)
Sweden dummy	0.173***	0.095*	0.094
	(0.062)	(0.057)	(0.068)
Norway dummy	0.245***	0.260***	0.262***
	(0.055)	(0.052)	(0.053)
Denmark dummy	0.032	0.101	0.101
	(0.092)	(0.092)	(0.093)
Primary schooling			−0.931**
			(0.386)
Secondary schooling	−1.142**		−0.949**
	(0.462)		(0.452)
Tertiary schooling		0.949***	
		(0.271)	
Agglomeration effect, Swe	−0.014	−0.014	−0.016
	(0.024)	(0.023)	(0.023)
Agglomeration effect, Nor	0.011	−0.005	−0.006
	(0.022)	(0.021)	(0.023)
Agglomeration effect, Den	0.014	−0.003	−0.007
	(0.025)	(0.024)	(0.026)
Constant	11.543***	10.800***	11.735***
	(0.221)	(0.072)	(0.227)
Observations	70	70	70
R^2	0.739	0.761	0.762
Adjusted R^2	0.705	0.730	0.726
F Statistic	21.592*** (df = 8; 61)	24.288*** (df = 8; 61)	21.354*** (df = 9; 60)
Note:			*p<0.1; **p<0.05; ***p<0.01

The results of the regressions are pretty clear in a statistical sense: there are no significant differences in agglomeration effects between any country and Finland. This effectively rejects our hypothesis of the existence of different size for effects in different countries. We can conclude that the results obtained in the previous section hold for all Nordic countries.

One interesting thing, however, that is visible also in this table is the statistical significance of the country dummy for Norway. This points once again to the fact that Norway has some characteristics that significantly affect the productivity positively. So, one thing learned from this sensitivity analysis is that Norway has a head start when it comes to productivity.

5. Conclusions and outlook

In this thesis we have been studying the presence and magnitude of agglomeration effect in the Nordic countries. It can be concluded that there exists a statistically significant positive agglomeration effect across all Nordic countries i.e. the agglomeration of employed population affects the productivity of the regions. The results of the OLS regressions show that this effect is present in all seven different model specifications and can therefore be considered to confirm the hypothesis of the existence of this effect.

Additionally, the magnitude of the agglomeration effect was found to be between 2.1 and 2.9 %. No evidence of differences in this magnitude between the Nordic countries was found in the alternative regression inspecting this possibility. The value for the magnitude of the effect is around 2-3 % smaller than in the corresponding research papers studying the same phenomenon in the five largest European countries and around 3-4 % smaller than in the USA. This result is as expected due to the unique geographical and political characteristics of the Nordic countries. Therefore, the results obtained in this study seem to be well in line with the existing literature.

The primary research method in this study has been linear OLS regression. While performing statistical testing, evidence of possible endogeneity of the explanatory variable was found. However, the IV estimators based on an instrument proposed by the existing literature, *log* of the total land area of NUTS 3 regions, yielded inconsistent estimates and could therefore not be considered as proper results. Due to this, and based on the very robust results across different model specifications obtained in the OLS regressions, it can be concluded that the OLS estimates probably yield fair estimates keeping the arisen endogeneity issue in mind.

Possibilities to further develop the research on this phenomenon in the Nordics would include to continue the work in finding a better instrument for the main explanatory variable. Also, the education of the employed population could be defined in even finer detail given there is plenty of data available. Some of the challenges of this topic are the low number of NUTS 3 regions in the Nordics that prevents the use of too complex model specifications and fairly large differences in geographical and demographical characteristics between regions. These point to the need to include some conservatism in the results to be able to be considered in all regions effectively. The doubling of employment density is, after all, very different thing to do in Helsinki than in Lapland.

In the end, the social scientific contribution of this thesis has been to bring more points of view to the political debate on the redistribution of investments and support between the growth centers and more sparsely populated regions in the Nordics. As both hypotheses were found to hold, it can be concluded that this objective was fulfilled. Now we know that it is also economically sensible to locate near each other, a finding that surely brings hope for the future.

Appendices

A. Region indices and data

Index	Region	GVA (M EUR)	Employment	Area (km^2)	Primary schooling ratio	Secondary schooling ratio	Tertiary schooling ratio
1	Keski-Suomi	7526.55	111020	16703	0.28	0.43	0.28
2	Etelä-Pohjanmaa	5051.18	86240	13444	0.32	0.44	0.24
3	Pohjanmaa	5976.75	86790	7753	0.30	0.40	0.29
4	Satakunta	6967.18	100840	7820	0.32	0.42	0.25
5	Pirkanmaa	15041.15	219520	12585	0.27	0.41	0.31
6	Uusimaa	70531.19	852940	9097	0.28	0.34	0.37
7	Varsinais-Suomi	13990.69	209080	10663	0.30	0.40	0.29
8	Kanta-Häme	4739.09	73990	5199	0.31	0.41	0.27
9	Päijät-Häme	5388.55	81220	5125	0.32	0.42	0.26
10	Kymenlaakso	5311.61	74080	5149	0.32	0.43	0.24
11	Etelä-Karjala	4141.39	52510	5327	0.32	0.42	0.25
12	Etelä-Savo	3807.39	60930	14257	0.32	0.43	0.24
13	Pohjois-Savo	6755.28	101510	16768	0.29	0.44	0.27
14	Pohjois-Karjala	4152.54	65600	17761	0.29	0.45	0.25
15	Kainuu	1821.3	33270	20197	0.31	0.45	0.23
16	Keski-Pohjanmaa	2091.4	31390	5020	0.32	0.43	0.24
17	Pohjois-Pohjanmaa	10784.5	166880	36815	0.27	0.44	0.28
18	Lappi	5476.09	76880	92674	0.29	0.45	0.25
19	Ahvenanmaa	1179.35	18330	1553	0.37	0.39	0.23
20	Stockholm	126918.16	1253000	6524	0.17	0.37	0.43
21	Uppsala	13458.28	165000	8190	0.19	0.40	0.40
22	Södermanland	8215.75	114000	6075	0.25	0.47	0.27
23	Östergötland	15766.61	207000	10559	0.22	0.44	0.32
24	Örebro	9880.15	138000	8504	0.23	0.46	0.29
25	Västermanland	9091.89	118000	5118	0.23	0.46	0.29
26	Jönköping	12143.48	178000	10437	0.26	0.46	0.27
27	Kronoberg	7177.74	95000	8424	0.24	0.45	0.29
28	Kalmar	7420.43	105000	11165	0.25	0.46	0.27
29	Gotland	1780.83	30000	3135	0.24	0.48	0.27
30	Blekinge	5004.54	68000	2931	0.24	0.45	0.29
31	Skåne	45308.71	597000	10968	0.21	0.41	0.35
32	Halland	9672.74	141000	5427	0.22	0.45	0.31
33	Västra Götaland	67346.13	817000	23880	0.22	0.43	0.34
34	Värmland	8587.69	117000	17519	0.22	0.48	0.28
35	Dalarna	9384.51	126000	29029	0.24	0.49	0.25
36	Gävleborg	8985.09	123000	18118	0.26	0.48	0.25
37	Västernorrland	8645.96	113000	21549	0.23	0.48	0.27
38	Jämtland	4078.47	59000	48935	0.21	0.49	0.28
39	Västerbotten	9084.94	126000	54665	0.18	0.46	0.35
40	Norrbotten	9643.56	120000	97239	0.20	0.51	0.28

Index	Region	GVA (M EUR)	Employment	Area (km^2)	Primary schooling ratio	Secondary schooling ratio	Tertiary schooling ratio
41	Oslo	55807.19	476000	454	0.21	0.30	0.48
42	Akershus	27182.89	278000	4918	0.24	0.38	0.36
43	Hedmark	7365.47	91000	27398	0.33	0.42	0.24
44	Oppland	7071.82	92000	25192	0.30	0.45	0.24
45	Østfold	10026.26	122000	4187	0.32	0.42	0.25
46	Buskerud	11758.29	130000	14912	0.29	0.42	0.29
47	Vestfold	9207.34	112000	2225	0.27	0.43	0.29
48	Telemark	6976.85	77000	15298	0.30	0.45	0.25
49	Aust-Agder	4172.81	50000	9155	0.28	0.44	0.28
50	Vest-Agder	8266.96	93000	7279	0.26	0.44	0.29
51	Rogaland	27598.66	271000	9377	0.26	0.42	0.31
52	Hordaland	27133.61	279000	15437	0.25	0.42	0.33
53	Sogn og Fjordane	4976.42	59000	18622	0.26	0.47	0.26
54	Møre og Romsdal	13036.45	139000	13985	0.28	0.45	0.26
55	Sør-Trøndelag	15449.29	166000	18848	0.24	0.41	0.34
56	Nord-Trøndelag	5088.94	65000	22412	0.28	0.46	0.25
57	Nordland	10307.05	120000	38478	0.32	0.42	0.25
58	Troms	7246.58	87000	25877	0.30	0.39	0.31
59	Finnmark	3289.31	41000	48631	0.35	0.38	0.26
60	Byen København	43469.64	490000	170	0.19	0.33	0.43
61	Københavns omegn	35636.10	334000	342	0.27	0.38	0.32
62	Nordsjælland	15449.07	183000	1449	0.25	0.38	0.34
63	Bornholm	1195.25	17000	592	0.34	0.44	0.20
64	Østsjælland	7117.73	100000	808	0.27	0.43	0.28
65	Vest- og Sydsjælland	16581.71	219000	6415	0.33	0.44	0.21
66	Fyn	15576.44	214000	3479	0.29	0.42	0.26
67	Sydjylland	29606.77	357000	8777	0.31	0.43	0.23
68	Vestjylland	30675.05	415000	5841	0.32	0.44	0.22
69	Østjylland	16525.53	218000	7154	0.26	0.41	0.31
70	Nordjylland	20204.46	274000	7879	0.30	0.43	0.24

Bibliography

Arzaghi, Mohammad and Vernon Henderson (2008). "Networking off madison avenue". In: *The Review of Economic Studies* 75.4, pp. 1011–1038.

Au, Chun-Chung and Vernon Henderson (2006). "Are Chinese cities too small?" In: *The Review of Economic Studies* 73.3, pp. 549–576.

Baum-Snow, Nathaniel and Daniel A Hartley (2016). "Accounting for central neighborhood change, 1980-2010". In: *FRB of Chicago Working Papers*. FRB of Chicago Working Paper No. WP-2016-9. Available at SSRN: https://ssrn.com/abstract=2839686.

Brülhart, Marius and Federica Sbergami (2009). "Agglomeration and growth: Cross-country evidence". In: *Journal of Urban Economics* 65, pp. 48–63.

Ciccone, Antonio (2002). "Agglomeration effects in Europe". In: *European Economic Review* 46, pp. 213–227.

Ciccone, Antonio and Robert E Hall (1996). "Productivity and the Density of Economic Activity". In: *American Economic Review* 86, pp. 54–70.

Datahub (2018). *European NUTS boundaries as GeoJSON at 1:60m*. GeoJSON data retrieved from Datahub.io database on 05.11.2018. URL: https://datahub.io/core/geo-nuts-administrative-boundaries.

Duranton, Gilles and Diego Puga (2004). "Micro-foundations of urban agglomeration economies". In: *Handbook of regional and urban economics*. Vol. 4. Elsevier, pp. 2063–2117.

Eurostat (2018a). *Area by NUTS 3 region*. Data retrieved from Eurostat database on 02.10.2018. URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=reg_area3&lang=en.

— (2018b). *Employment (thousand persons) by NUTS 3 regions*. Data retrieved from Eurostat database on 02.10.2018. URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3empers&lang=en&fbclid=IwAR0E-LuXcsxgIBUaD38sY8aHN-6Wh9IKcdLpzx-ccqaf5nnfJ7sVTwTE9GQ.

— (2018c). *Gross value added at basic prices by NUTS 3 regions*. Data retrieved from Eurostat database on 02.10.2018. URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3gva&lang=en&fbclid=IwAR3HpzkmV1YVNj3cNnnjlyHnEBc_-6ey3uCVmqoQ4Xj-33Nq-4wiEixgMsc.

Henderson, Vernon (2003a). "Marshall's scale economies". In: *Journal of urban economics* 53.1, pp. 1–28.

— (2003b). "The urbanization process and economic growth: The so-what question". In: *Journal of Economic growth* 8.1, pp. 47–71.

Krugman, Paul (1991). "Increasing returns and economic geography". In: *Journal of political economy* 99.3, pp. 483–499.

Marshall, Alfred (1920). "Principles of Economics (London, 1920)". In: *Book VI*, pp. 618–619.

Moretti, Enrico (2004). "Human capital externalities in cities". In: *Handbook of regional and urban economics* 4, pp. 2243–2291.

OECD (2019). *Value added by activity (indicator)*. Accessed on 11 April 2019. URL: 10.1787/a8b2bd2b-en.

Pudding (2018). *Human Terrain - Visualizing the World's Population in 3D*. Accessed on 05.04.2019. URL: https://pudding.cool/2018/10/city_3d/.

Quigley, John M. (2008). "Urbanization, agglomeration and economic development". In: *World Bank Commission on Growth and Development series of working papers*. URL: http://siteresources.worldbank.org/EXTPREMNET/Resources/%20%5C%5C%09%09%09%09%09489960-1338997241035/Growth_Commission_Working_Paper_19_Urbanization_%20%5C%5C%20%09%09%09%09%09Agglomeration_Economic_Development.pdf.

Ricardo, David (1817). *On the Principles of Political Economy and Taxation*. London.

Rosenthal, Stuart S and William C Strange (2004). "Evidence on the nature and sources of agglomeration economies". In: *Handbook of regional and urban economics*. Vol. 4. Elsevier, pp. 2119–2171.

Smith, Adam (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Vol. 1. London.

Statistics Denmark (2018). *Educational attainment (15-69 years) by region, ancestry, highest education completed, age and sex*. Data retrieved from the Statistics Denmark database on 02.10.2018. URL: http://www.statbank.dk/10139.

Statistics Finland (2018). *Suomen virallinen tilasto (SVT): Väestön koulutusrakenne [verkkojulkaisu]*. ISSN=1799-4586. Helsinki. Data retrieved from Statistics Finland database on 02.10.2018. URL: http://www.stat.fi/til/vkour/index.html?fbclid=IwAR3GnAnyMAy05AVIgAQMx99MHGRWmjZ6chJcpwZUxslbIMPBOapq5GRvjeI.

Statistics Norway (2018). *Educational attainment, by country, age and sex (C) 1980-2017*. Data retrieved from the Statistics Norway database on 02.10.2018. URL: https://www.ssb.no/en/statbank/table/08921/.

Statistics Sweden (2018). *The population 1990-2017 by level of education and sex. Age 25-64*. Data retrieved from the Statistics Sweden database on 02.10.2018. URL: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/education-and-research/education-of-the-population/educational-attainment-of-the-population/.

Stewart, Thomas and Clare Ruckdeschel (1998). "Intellectual capital: The new wealth of organizations". In: *Performance Improvement* 37.7, pp. 56–59.

The Child and Youth Administration, City of Copenhagen (2019). *How is the Danish school system structured?* Accessed on 11 April 2019. URL: https://international.kk.dk/artikel/how-danish-school-system-structured.

Williamson, Jeffrey G (1965). "Regional inequality and the process of national development: a description of the patterns". In: *Economic development and cultural change* 13.4, Part 2, pp. 1–84.